

The Closest Pair

1 second, 2 megabytes

Hierarchical Clustering เป็นเทคนิคการจัดกลุ่มที่สร้างโครงสร้างลำดับชั้นระหว่างกลุ่ม (Clusters) โดยมีสองวิธีหลัก ๆ คือ

- Agglomerative (Bottom-up): เริ่มต้นจากการพิจารณาจุดข้อมูลแต่ละจุดเป็นคลัสเตอร์แยกกัน จากนั้นรวมคลัสเตอร์ที่ใกล้ที่สุดเข้าด้วยกันทีละคู่จนกว่าจะได้คลัสเตอร์เดียว
- Divisive (Top-down): เริ่มจากคลัสเตอร์ทั้งหมดแล้วแบ่งย่อยออกไปตามระยะห่าง

ถ้าหากนิสิตจะต้องออกแบบขั้นตอนวิธีในการหาระยะทางที่สั้นที่สุด เพื่อใช้ในการทำ Hierarchical Clustering แบบ Agglomerative (Bottom-up) โดยจะต้องรวมคลัสเตอร์ที่ใกล้ที่สุดเข้าด้วยกันทีละคู่จนกว่าจะได้คลัสเตอร์เดียว

ทั้งนี้ เทคนิคอาจถูกนำไปใช้วิเคราะห์ขั้นตอนวิธีการเรียนรู้ของเครื่อง (Machine Learning Algorithm)

โจทย์ จงเขียนโปรแกรมรับข้อมูลจุดแต่ละจุดบนระนาบสองมิติ เพื่อแสดงระยะทางของคู่จุดที่สั้นที่สุด

ข้อมูลนำเข้า (Input)

บรรทัดแรก รับจำนวนชุดทดสอบย่อย t ชุด โดยที่ $1 \leq t \leq 20$

บรรทัดถัดไปจะรับข้อมูลของแต่ละชุดทดสอบจำนวน t ชุด ซึ่งแต่ละชุดมีดังนี้

บรรทัดแรก รับจำนวนจุดที่อยู่ในระนาบสองมิติทั้งหมดของทั้งเส้น n_i จุด

โดยที่ $2 \leq n \leq 20000$

บรรทัดที่สอง ถึงบรรทัดที่ $n + 1$

รับพิกัด (x, y) ของแต่ละจุด โดยที่ $x, y \in \mathbb{R}^+$ และ $x, y \leq 10^5$

ข้อมูลส่งออก (Output)

มีหลายบรรทัด แสดงระยะทางที่สั้นที่สุดเป็นทศนิยม 6 ตำแหน่ง ของคู่จุดบนระนาบสองมิติที่เป็นไปได้ของแต่ละ t

ตัวอย่างข้อมูลนำเข้าและส่งออก (Input/Output Example)

ตัวอย่างข้อมูลนำเข้า	ตัวอย่างข้อมูลส่งออก
1 5 (3.14, 2.71) (1.23, 4.56) (7.89, 0.12) (5.67, 8.90) (9.01, 1.23)	1.576864
2 3 (1, 0) (0, 1) (0.5, 4) 5 (1.745, 2.228) (4.355, 1.772) (2.50, 3.51) (3.298, 2.84) (3.0, 3.0)	1.414214 0.338237

เกณฑ์การให้คะแนนและขอบเขตปัญหาย่อย (Scoring Criterias / Subproblems)

ให้คะแนนตามจำนวนชุดทดสอบที่ทำสำเร็จ

ระดับ	ปัญหาย่อย	Runtime และ Memory	ชุดทดสอบ	คะแนน
1	$n \leq 500$	1 second, 2 megabytes	7 ชุด	50%
2	$n \leq 10000$	0.75 seconds, 2 megabytes	12 ชุด	70%
3	ไม่มีเงื่อนไขเพิ่มเติม	0.5 seconds, 2 megabytes	16 ชุด	100%